STIC EIC 2100 Search Request Form

10847

Today's Date: 11/18/03

What date would you like to use to limit the search?
Priority Date: 4/18/00 Other:

Name Nathan Hilleoy
AU 2176 Examiner # 79871
Room # 3C19 Phone 5-4502
Serial # 09/551,402

Format for Search Results (Circle One):
PAPER DISK EMAIL

Where have you searched so far?
USP DWPI EPO JPO ACM IBM TDB
IEEE INSPEC SPI Other _____

Is this a "Fast & Focused" Search Request? (Circle One) YES NO
A "Fast & Focused" Search is completed in 2-3 hours (maximum). The search must be on a very specific topic and meet certain criteria. The criteria are posted in EIC2100 and on the EIC2100 NPL Web Page at http://ptoweb/patents/stic/stic-tc2100.htm.

What is the topic, novelty, motivation, utility, or other specific details defining the desired focus of this search? Please include the concepts, synonyms, keywords, acronyms, definitions, strategies, and anything else that helps to describe the topic. Please attach a copy of the abstract, background, brief summary, pertinent claims and any citations of relevant art you have found.

Claims: 30-38

STIC Searcher Geoffrey St. Leger Phone 308-7800
Date picked up 11/18/3 Date Completed 11/18/3

sira Search and Information Resources Administration

STIC Search Report
EIC 2100

STIC Database Tracking Number: 108471

TO:	From:
Nathan Hillery **Location:** **Art Unit : 2176** **Tuesday, November 18, 2003** **Case Serial Number:**	**Geoffrey St. Leger** **Location: EIC 2100** **PK2-4B30** **Phone: 308-7800** **geoffrey.stleger@uspto.gov**

Search Notes

Dear Examiner Hillery,

Attached please find the results of your Fast & Focused search request for application . I searched Dialog's foreign patent files, technical databases, product announcement files and general files; along with the Internet.

Please let me know if you have any questions.

Regards,

Geoffrey St. Leger
4B30/308-7800

STIC Search Results Feedback Form

USPTO

EIC 2100

Questions about the scope or the results of the search? Contact ***the EIC searcher or contact***:

Anne Hendrickson, EIC 2100 Team Leader
308-7831, CPK2-4B40

Voluntary Results Feedback Form

- *I am an examiner in Workgroup:* ☐ *Example: 3730*
- *Relevant prior art **found**, search results used as follows:*
  - ☐ 102 rejection
  - ☐ 103 rejection
  - ☐ Cited as being of interest.
  - ☐ Helped examiner better understand the invention.
  - ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
  (journal articles, conference proceedings, new product announcements etc.)

- *Relevant prior art **not found**:*
  - ☐ Results verified the lack of relevant prior art (helped determine patentability).
  - ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC/EIC2100 CPK2-4B40

sira
Search and Information Resources Administration

```
File 347:JAPIO Oct 1976-200[redacted]Jul(Updated 031105)
         (c) 2003 JPO & JAPIO
File 350:Derwent WPIX 1963-2003/UD,UM &UP=200374
         (c) 2003 Thomson Derwent

Set     Items   Description
S1       1697   XML OR (EXTENSIBLE OR XTENSIBLE)()(MARKUP OR MARK()UP)
S2        763   NAME? ?(3N)ATTRIBUTE? ?
S3       2394   VALUE? ?(3N)ATTRIBUTE? ?
S4         25   S2(5N)S3(5N)(PAIR??? OR COUPLE? ? OR SET? ? OR GROUP???? OR
                TOKEN? ?)
S5     344083   TABLE? ? OR TABLESPACE? ? OR LUT? ? OR HASHTABLE? ?
S6        111   CODESPACE? ? OR CODE()SPACE? ?
S7          1   S1 AND S4
S8          0   S1 AND S6
S9          6   S1 AND S2 AND S3
S10         6   S7 OR S9
```

10/5/1 (Item 1 from file: 347)
DIALOG(R)File 347:JAPIO

06364595 **Image available**
DOCUMENT FILE RETRIEVAL DEVICE AND MACHINE READABLE RECORDING MEDIUM RECORDING PROGRAM

PUB. NO.: 11-306205 [JP 11306205 A]
PUBLISHED: November 05, 1999 (19991105)
INVENTOR(s): SHIMAZU HIDEO
APPLICANT(s): NEC CORP
APPL. NO.: 10-129485 [JP 98129485]
FILED: April 23, 1998 (19980423)
INTL CLASS: G06F-017/30

ABSTRACT

PROBLEM TO BE SOLVED: To realize a retrieval inquiry about a WWW home page by a natural language.

SOLUTION: A WWW home page being a retrieval object document file is described in an **XML** . When a retrieval condition composition is inputted, a keyword extraction part 4 converts a natural language expression expressing an **attribute** name into an attribute name index including the attribute name and also converts the natural language expression expressing the attribute value into an **attribute value** index including a pair of the said **attribute name** and **attribute value** . A keyword filter part 5 deletes the **attribute name** index existing at a place where the **attribute name** and the **attribute value** of the same **attribute** exist adjacent to each other in a converted index string. A document contents check part 6 checks whether or not a tag corresponding to **pairs** of **attribute name** and **value** of the all **attribute value** index exists in the retrieval object document file. If the said tag exists, a document contents output part 9 retrieves and outputs the **attribute value** of the tag having the relevant **attribute name** of the **attribute name** index that is included in the converted index string.

10/5/2 (Item 1 from file: 350)
DIALOG(R)File 350:Derwent WPIX

015589150 **Image available**
WPI Acc No: 2003-651305/200362
XRPX Acc No: N03-518254
XML **document conversion method involves converting tag name of non-key component into** attribute value **corresponding to prescribed** attribute name **assigned to new component**
Patent Assignee: FUJITSU LTD (FUIT)
Inventor: ITANI N; YAHAGI H; YOSHIDA S
Number of Countries: 002 Number of Patents: 002
Patent Family:

Patent No	Kind	Date	Applicat No	Kind	Date	Week	
JP 2003203067	A	20030718	JP 2001401934	A	20011228	200362	B
US 20030158854	A1	20030821	US 2002274230	A	20021021	200362	

Priority Applications (No Type Date): JP 2001401934 A 20011228
Patent Details:

Patent No	Kind	Lan	Pg	Main IPC	Filing Notes
JP 2003203067	A		56	G06F-017/21	
US 20030158854	A1			G06F-007/00	

Abstract (Basic): JP 2003203067 A

NOVELTY - A tag name assigned to a non-key component of an **XML** document, is converted into an **attribute** value corresponding to prescribed **attribute name** which is assigned to a new component. The

character string in the tag name of the non-key component is defined as the content of the new component, and the key component of the **XML** document is retained.

DETAILED DESCRIPTION - An INDEPENDENT CLAIM is also included for data conversion method.

USE - For converting character strings of **XML** document.

ADVANTAGE - The **XML** documents are converted into compressed data values, thereby reducing the memory space for storing the **XML** documents. Also the processing line is reduced.

DESCRIPTION OF DRAWING(S) - The figure shows the block diagram of the data structure conversion system. (Drawing includes non-English language text).

XML document conversion processor (10)

extensible style sheet language transformation (XSLT) converter (11)

XSLT structural transformation unit (12)

application software (30)

pp; 56 DwgNo 2/48

Title Terms: DOCUMENT; CONVERT; METHOD; CONVERT; TAG; NAME; NON; KEY; COMPONENT; ATTRIBUTE; VALUE; CORRESPOND; PRESCRIBED; ATTRIBUTE; NAME; ASSIGN; NEW; COMPONENT

Derwent Class: T01

International Patent Class (Main): G06F-007/00; G06F-017/21

International Patent Class (Additional): G06F-012/00

File Segment: EPI

10/5/3 (Item 2 from file: 350)

DIALOG(R)File 350:Derwent WPIX

014385694 **Image available**

WPI Acc No: 2002-206397/200226

XRPX Acc No: N02-157188

Matching source and target data by comparing source and target data nodes to determine percentage measure of similarity

Patent Assignee: INFOGLIDE CORP (INFO-N); RIPLEY J R (RIPL-I); WHEELER D B (WHEE-I); WOTRING S C (WOTR-I)

Inventor: RIPLEY J R; WHEELER D B; WOTRING S C

Number of Countries: 097 Number of Patents: 004

Patent Family:

Patent No	Kind	Date	Applicat No	Kind	Date	Week	
WO 200213049	A1	20020214	WO 2001US24628	A	20010806	200226	B
US 20020055932	A1	20020509	US 2000223449	P	20000804	200235	
			US 2001682207	A	20010806		
AU 200181111	A	20020218	AU 200181111	A	20010806	200244	
EP 1317715	A1	20030611	EP 2001959570	A	20010806	200339	
			WO 2001US24628	A	20010806		

Priority Applications (No Type Date): US 2000223449 P 20000804; US 2001682207 A 20010806

Patent Details:

Patent No	Kind	Lan	Pg	Main IPC	Filing Notes
WO 200213049	A1	E	64	G06F-017/00	

Designated States (National): AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CO CR CU CZ DE DK DM DZ EC EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR LS LT LU LV MA MD MG MK MN MW MX MZ NO NZ PL PT RO RU SD SE SG SI SK SL TJ TM TR TT TZ UA UG UZ VN YU ZA ZW

Designated States (Regional): AT BE CH CY DE DK EA ES FI FR GB GH GM GR IE IT KE LS LU MC MW MZ NL OA PT SD SE SL SZ TR TZ UG ZW

US 20020055932	A1		G06F-007/00	Provisional application US 2000223449
AU 200181111	A		G06F-017/00	Based on patent WO 200213049
EP 1317715	A1	E	G06F-017/00	Based on patent WO 200213049

Designated States (Regional): AL AT BE CH CY DE DK ES FI FR GB GR IE IT LI LT LU LV MC MK NL PT RO SE SI TR

Abstract (Basic): WO 200213049 A1

NOVELTY - Method consists in selecting comparison methods such as exact string match, similarity string composition, synonym table lookup etc., comparing the source and target data structure nodes and determining a measure of similarity between them. Each node comprises an element **name** , data type **attribute** and an **attribute** description **value** .

DETAILED DESCRIPTION - A strategy list assigns the comparison methods to each node name and value and data is automatically mapped to the target data node if the similarity measure exceeds a threshold. Each node is represented by HTML, **XML** or SGML and the comparison steps are repeated recursively.

There is an INDEPENDENT CLAIM for a data matching computer program.

USE - Method is for sharing data held in different databases with different formats and structures over the Internet

ADVANTAGE - Method saves time and money by not requiring data sources to homogenize information before interchanging.

DESCRIPTION OF DRAWING(S) - The figure shows an overview of heterogeneous database searching.

pp; 64 DwgNo 1/11

Title Terms: MATCH; SOURCE; TARGET; DATA; COMPARE; SOURCE; TARGET; DATA; NODE; DETERMINE; PERCENTAGE; MEASURE; SIMILAR

Derwent Class: T01

International Patent Class (Main): G06F-007/00; G06F-017/00

International Patent Class (Additional): G06F-017/30

File Segment: EPI

10/5/4 (Item 3 from file: 350)

DIALOG(R)File 350:Derwent WPIX

013827738 **Image available**

WPI Acc No: 2001-311950/200133

XRPX Acc No: N01-223668

Attribute extractor for structurized documents, extracts and outputs attribute value corresponding to indexed position, obtained by comparing input document content with prestored attribute schema

Patent Assignee: FUJI XEROX CO LTD (XERF)

Number of Countries: 001 Number of Patents: 001

Patent Family:

Patent No	Kind	Date	Applicat No	Kind	Date	Week	
JP 2001075974	A	20010323	JP 99246880	A	19990901	200133	B

Priority Applications (No Type Date): JP 99246880 A 19990901

Patent Details:

Patent No	Kind	Lan	Pg	Main IPC	Filing Notes
JP 2001075974	A		25	G06F-017/30	

Abstract (Basic): JP 2001075974 A

NOVELTY - Contents of input document (1a) are compared with prestored **attribute** schema (1f). **Attribute name** and its index position corresponding to document content are extracted from attribute schema, respectively by extractors (1b,1c). **Attribute names** for position not indexed are deleted. Attribute data corresponding to indexed positions is extracted by extractor (1d) and outputs the data as a list (1e).

DETAILED DESCRIPTION - INDEPENDENT CLAIMS are also included for the following:

(a) Attribute extracting method;

(b) Recording medium with attribute extracting program

USE - For detecting convergence of attributes in structurized documents specified in standard generalized markup language (SGML), **extensible markup** language (**XML**).

ADVANTAGE - Required attribute is extracted simply without breaking the format of document and without being conscious of variations in document.

DESCRIPTION OF DRAWING(S) - The figure shows the conceptional

diagram of attribute extractor (The drawing includes non-English language text).
Input document (1a)
Extractors (1b-1d)
List (1e)
Attribute schema (1f)
pp; 25 DwgNo 1/29
Title Terms: ATTRIBUTE; EXTRACT; DOCUMENT; EXTRACT; OUTPUT; ATTRIBUTE; VALUE; CORRESPOND; INDEX; POSITION; OBTAIN; COMPARE; INPUT; DOCUMENT; CONTENT; ATTRIBUTE
Derwent Class: T01
International Patent Class (Main): G06F-017/30
International Patent Class (Additional): G06F-017/21; G06F-017/27
File Segment: EPI

10/5/5 (Item 4 from file: 350)
DIALOG(R)File 350:Derwent WPIX

013461666 **Image available**
WPI Acc No: 2000-633609/200061
XRPX Acc No: N00-469628
Attribute extractor of structurizing document, compares structurizing document with character row pattern, based on which attribute name **and** attribute value **of structurizing document are extracted**
Patent Assignee: FUJI XEROX CO LTD (XERF)
Number of Countries: 001 Number of Patents: 001
Patent Family:

Patent No	Kind	Date	Applicat No	Kind	Date	Week	
JP 2000259660	A	20000922	JP 9964504	A	19990311	200061	B

Priority Applications (No Type Date): JP 9964504 A 19990311
Patent Details:

Patent No	Kind	Lan	Pg	Main IPC	Filing Notes
JP 2000259660	A		21	G06F-017/30	

Abstract (Basic): JP 2000259660 A
NOVELTY - **Attribute name** showing the **attribute** of a structurizing document (1a) and character row pattern corresponding to the **attribute name** , are defined by a schema definition unit (1b). The structurizing document is compared with the character row pattern, based on which **attribute name** and **attribute value** of the structurizing document are extracted.
USE - For extracting and grouping row of desired attribute from structurizing document such as hypertext markup language document, **extensible markup** language document, standard generalized mark up language document.
ADVANTAGE - Enables extracting required **attribute name** and **attribute value** . Enables identifying paragraph between documents.
DESCRIPTION OF DRAWING(S) - The figure shows the block diagram of attribute extractor.
Structurizing document (1a)
Schema definition unit (1b)
pp; 21 DwgNo 1/39
Title Terms: ATTRIBUTE; EXTRACT; DOCUMENT; COMPARE; DOCUMENT; CHARACTER; ROW; PATTERN; BASED; ATTRIBUTE; NAME; ATTRIBUTE; VALUE; DOCUMENT; EXTRACT
Derwent Class: T01
International Patent Class (Main): G06F-017/30
International Patent Class (Additional): G06F-017/21; G06F-017/27
File Segment: EPI

10/5/6 (Item 5 from file: 350)
DIALOG(R)File 350:Derwent WPIX

012872807 **Image available**

WPI Acc No: 2000-044640/200[redacted]
XRPX Acc No: N00-034220

Text file searching system in internet - has keyword filter to selectively delete attribute name **index when its repetition is detected**

Patent Assignee: NEC CORP (NIDE)
Number of Countries: 001 Number of Patents: 002
Patent Family:

Patent No	Kind	Date	Applicat No	Kind	Date	Week	
JP 11306205	A	19991105	JP 98129485	A	19980423	200004	B
JP 3191762	B2	20010723	JP 98129485	A	19980423	200143	

Priority Applications (No Type Date): JP 98129485 A 19980423
Patent Details:

Patent No	Kind	Lan	Pg	Main IPC	Filing Notes
JP 11306205	A		21	G06F-017/30	
JP 3191762	B2		21	G06F-017/30	Previous Publ. patent JP 11306205

Abstract (Basic): JP 11306205 A

NOVELTY - File search demand in natural language expression is investigated to acquire **attribute** name index and **attribute value** , using which a keyword for searching, is extracted. A filter (5) selectively deletes **attribute name** index when its repetition is detected. **Attribute value** and **name** index from filter are then used to search the required file. DETAILED DESCRIPTION - An INDEPENDENT CLAIM is also included for recording medium storing text file searching program.

USE - For searching text file such as **XML** in internet by natural language expression search inquiry.

ADVANTAGE - Redundancy of reply corresponding to search demand is eliminated by keyword filter. User desired file can be retrieved easily, by natural language expression demand. DESCRIPTION OF DRAWING(S) - The figure shows the block diagram of text file searching system. (5) Keyword filter.

Dwg.1/7

Title Terms: TEXT; FILE; SEARCH; SYSTEM; KEYWORD; FILTER; SELECT; DELETE; ATTRIBUTE; NAME; INDEX; REPEAT; DETECT
Derwent Class: T01; W01
International Patent Class (Main): G06F-017/30
File Segment: EPI

File 348:EUROPEAN PATENTS 1978-2003/Nov W02
(c) 2003 European Patent Office
File 349:PCT FULLTEXT 1979-2002/UB=20031113,UT=20031106
(c) 2003 WIPO/Univentio

Set	Items	Description
S1	6734	XML OR (EXTENSIBLE OR XTENSIBLE)()(MARKUP OR MARK()UP)
S2	2755	NAME? ?(3N)ATTRIBUTE? ?
S3	4920	VALUE? ?(3N)ATTRIBUTE? ?
S4	261	S2(5N)S3(5N)(PAIR??? OR COUPLE? ? OR SET? ? OR GROUP???? OR TOKEN? ?)
S5	461192	TABLE? ? OR TABLESPACE? ? OR LUT? ? OR HASHTABLE? ?
S6	574	CODESPACE? ? OR CODE()SPACE? ?
S7	53	S1(S)S4 OR S1(100N)S4
S8	7	S1(S)S4(S)S5:S6
S9	12	S1(100N)S4(100N)S5:S6
S10	35	S1(S)S2(S)S3(S)S5:S6
S11	42	S8:S10

11/5,K/7 (Item 7 from file: 348)
DIALOG(R)File 348:EUROPEAN PATENTS

01334682
Method and computer system for publishing information
Verfahren und System zum Publizieren von Informationen
Methode et systeme pour la publication d'informations
PATENT ASSIGNEE:
iUniverse. com, Inc., (3276360), 910 E. Hamilton Avenue, Suite 100, Campbell, CA 95008, (US), (Applicant designated States: all)
INVENTOR:
Tam, Richard K., 15498 Via Caballero, Monte Sereno, CA 95030, (US)
Dunbar, Steve M., 900 Pepper Tree Lane, No. 1718, Santa Clara, CA 95051, (US)
Nguyen, Young C., 3238 Via Del Mar, San Jose, CA 95124, (US)
LEGAL REPRESENTATIVE:
Kirschner, Klaus Dieter, Dipl.-Phys. (6506), Schneiders & Behrendt Rechtsanwalte - Patentanwalte Sollner Strasse 38, 81479 Munchen, (DE)
PATENT (CC, No, Kind, Date): EP 1139253 A1 011004 (Basic)
APPLICATION (CC, No, Date): EP 2001106127 010313;
PRIORITY (CC, No, Date): US 536192 000326
DESIGNATED STATES: DE; FR; GB; IT
EXTENDED DESIGNATED STATES: AL; LT; LV; MK; RO; SI
INTERNATIONAL PATENT CLASS: G06F-017/60

ABSTRACT EP 1139253 A1
A method and a system take submissions of information offered for distribution or sale, combines partially or entirely at least two submissions to form a combination and distributing the combination in one or more forms. The system takes submissions from authors automatically over a network such as the Internet. Authors provide files for publication e.g., in XML files. Authors also provide the contractual terms for their publications. The system stores the submissions in two parts: content and descriptors that describe the content. On receipt of an order for distribution or sales, the system combines the contents and descriptors from at least two submissions to form a combination of the submissions. The system stores the contractual terms so that the authors are paid according to the distribution or sales of his or her publications. Customers provide their purchase orders for publications to the system. Customers can purchase publications in part or in whole as permitted by the contractual terms regarding the publications. Customers can also combine a publication with another publication or a personalized content submitted by the customers as permitted by the contractual terms regarding the publications. Customers can further select the output forms of their purchases, such as print media or electronic media, as permitted by the contractual terms regarding the publications.
ABSTRACT WORD COUNT: 214
NOTE:
Figure number on first page: 1

LEGAL STATUS (Type, Pub Date, Kind, Text):
Application: 011004 A1 Published application with search report
Examination: 020612 A1 Date of request for examination: 20020403
Change: 020821 A1 Designated contracting states changed 20020628
Withdrawal: 030723 A1 Date of withdrawal of application: 20030526
LANGUAGE (Publication,Procedural,Application): English; English; English
FULLTEXT AVAILABILITY:

Available Text	Language	Update	Word Count
CLAIMS A	(English)	200140	1113
SPEC A	(English)	200140	13362
Total word count - document A			14475
Total word count - document B			0
Total word count - documents A + B			14475

...SPECIFICATION is performed by software shown in the middle of p. 7 in file Parse.sqlj of Appendix D. In action 44, ingest engine 29 saves **attribute names** and **attribute values** of the **XML** tag node in

iu(underscore)attribute table 167 of content management database 32. Action 44 is followed by action 47. Action 44 is performed by software shown in the middle of p...

11/5,K/8 (Item 8 from file: 348)

DIALOG(R)File 348:EUROPEAN PATENTS

01286192

A system for treating saved queries as searchable documents in a document management system

System zum Behandeln von abgespeicherten Suchanfragen als durchsuchbare Dokumente in einem Dokumentenmanagementsystem

Systeme de traitement d'interrogations stockees comme documents qu'on peut chercher dans un systeme de gestion de documents

PATENT ASSIGNEE:
Ricoh Company, Ltd., (209037), 3-6, Nakamagome 1-chome, Ohta-ku, Tokyo 143-8555, (JP), (Proprietor designated states: all)
INVENTOR:
Persol, Kurt, 2882 Sand Hill Road, Suite 115, Menlo Park, CA 94025-7022, (US)
LEGAL REPRESENTATIVE:
Schwabe - Sandmair - Marx (100951), Stuntzstrasse 16, 81677 Munchen, (DE)

PATENT (CC, No, Kind, Date): EP 1104901 A2 010606 (Basic)
EP 1104901 A3 011107
EP 1104901 B1 030205
APPLICATION (CC, No, Date): EP 2000117365 000823;
PRIORITY (CC, No, Date): US 439151 991112
DESIGNATED STATES: DE; ES; FR; GB; IT; NL
EXTENDED DESIGNATED STATES: AL; LT; LV; MK; RO; SI
INTERNATIONAL PATENT CLASS: G06F-017/30
CITED PATENTS (EP B): EP 886228 A; US 5157783 A; US 5802512 A

ABSTRACT EP 1104901 A2

A file management appliance ("FMA") is a device that utilizes multiple processes and queues to provide document capture and indexing services as part of a document management system. Through the document capture and indexing services of an FMA-based system, documents are archived into one or more data storage devices, thereby forming a document database. One mechanism by which users may search for and access such archived data from one or more document databases is by formulating and submitting one or more queries to the FMA system. Queries formulated within an FMA system are treated as documents within the FMA system and accordingly may be archived within the document database for later retrieval and execution.

ABSTRACT WORD COUNT: 115
NOTE:
Figure number on first page: 2

LEGAL STATUS (Type, Pub Date, Kind, Text):

Application:	010606 A2	Published application without search report
Examination:	010606 A2	Date of request for examination: 20000823
Search Report:	011107 A3	Separate publication of the search report
Examination:	020508 A2	Date of dispatch of the first examination report: 20020402
Grant:	030205 B1	Granted patent

LANGUAGE (Publication,Procedural,Application): English; English; English
FULLTEXT AVAILABILITY:

Available Text	Language	Update	Word Count
CLAIMS A	(English)	200123	843
CLAIMS B	(English)	200306	843
CLAIMS B	(German)	200306	850
CLAIMS B	(French)	200306	1447
SPEC A	(English)	200123	7530
SPEC B	(English)	200306	7549
Total word count - document A			8374
Total word count - document B			10689

Total word count - document A + B 19063

...SPECIFICATION file contains special information about the document such as, for example, bibliographic data extracted from the capturing device. In one embodiment, document metadata consists of **pairs** of **attribute names** and their **values** .

Figure 4 is a **table** illustrating one embodiment of an FMA metadata file. In Figure 4, document metadata attributes are listed along with each attribute's meaning.

Figure 5 is a **table** illustrating a second embodiment of an FMA metadata file. In Figure 5, document metadata attributes are listed along with their acceptable value types.

Figure 6 illustrates one embodiment of an FMA metadata file in **extensible markup** language (**XML**). The partial metadata code depicted in Figure 6 is illustrative of what might be produced for a document that was captured by user "jones" (line...

...up to disk 37 (line 690).

In the event that an FMA encounters a metadata file that is not well-formed (as defined by the **XML** specification available from the World Wide Web Consortium (W3C) at http://www.w3.org), then in one embodiment, that FMA replaces the metadata with a...

...SPECIFICATION file contains special information about the document such as, for example, bibliographic data extracted from the capturing device. In one embodiment, document metadata consists of **pairs** of **attribute names** and their **values** .

Figure 4 is a **table** illustrating one embodiment of an FMA metadata file. In Figure 4, document metadata attributes are listed along with each attribute's meaning.

Figure 5 is a **table** illustrating a second embodiment of ...In Figure 5, document metadata attributes are listed along with their acceptable value types.

Figure 6 illustrates one embodiment of an FMA metadata file in **extensible markup** language (**XML**). The partial metadata code depicted in Figure 6 is illustrative of what might be produced for a document that was captured by user "jones" (line...

...up to disk 37 (line 690).

In the event that an FMA encounters a metadata file that is not well-formed (as defined by the **XML** specification available from the World Wide Web Consortium (W3C) at http://www.w3.org), then in one embodiment, that FMA replaces the metadata with a...

11/5,K/37 (Item 28 from file: 349)

00843106 **Image available**

SYSTEM AND METHOD FOR ESTABLISHING ELECTRONIC BUSINESS SYSTEMS FOR SUPPORTING COMMUNICATIONS SERVICES COMMERCE

SYSTEME ET PROCEDE PERMETTANT D'ETABLIR DES SYSTEMES DE COMMERCE ELECTRONIQUE POUR LE SUPPORT DU COMMERCE PAR DES SERVICES DE COMMUNICATION

Patent Applicant/Assignee:
CYGENT INC, 201 3rd Street, 2nd Floor, San Francisco, CA 94103, US, US (Residence), US (Nationality), (For all designated states except: US)

Patent Applicant/Inventor:
ELMORE Kevin, 653 Park Hill Road, Danville, CA 94526, US, US (Residence), US (Nationality), (Designated only for: US)
HOESER Vince, 20 Beaumont Court, Lafayette, CA 94549, US, US (Residence), US (Nationality), (Designated only for: US)
HOM David, 187 Magellan Avenue, San Francisco, CA 94116, US, US (Residence), US (Nationality), (Designated only for: US)
HORNE Linda, 6026 Ridgemont Drive, Oakland, CA 94619, US, US (Residence), US (Nationality), (Designated only for: US)
KAAKE Charles, 1931 Filbert Street, San Francisco, CA 94123, US, US

(Residence), US (Nationality), (Designated only for: US)
KELLY Kevin, 2334 Divisadero Street, #4, San Francisco, CA 94115, US, US
(Residence), US (Nationality), (Designated only for: US)
MABEL Mark, 796 Ashbury Street, San Francisco, CA 94117, US, US
(Residence), US (Nationality), (Designated only for: US)
RODRIGUEZ Hal, 619 25th Avenue #2, San Francisco, CA 94121, US, US
(Residence), US (Nationality), (Designated only for: US)
STEPHENS Paul, 957 Oak Street, #C, San Francisco, CA 94117, US, US
(Residence), US (Nationality), (Designated only for: US)
POKOTYLO Vadim, 20008 Gem Court, Castro Valley, CA 94546, US, US
(Residence), MD (Nationality), (Designated only for: US)
BAKHRU Girish, 225 Irving Street #9, San Francisco, CA 94122, US, US
(Residence), IN (Nationality), (Designated only for: US)
Legal Representative:
GLENN Michael (et al) (agent), Glenn Patent Group, Suite L., 3475 Edison Way, Menlo Park, CA 94025, US,
Patent and Priority Information (Country, Number, Date):
Patent: WO 200175549 A2-A3 20011011 (WO 0175549)
Application: WO 2001US10473 20010330 (PCT/WO US0110473)
Priority Application: US 2000193315 20000330
Designated States: AE AL AM AT AU AZ BA BB BG BR BY CA CH CN CR CU CZ DE DK DM EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR LS LT LU LV MA MD MG MK MN MW MX NO NZ PL PT RO RU SD SE SG SI SK SL TJ TM TR TT TZ UA UG US UZ VN YU ZA ZW
(EP) AT BE CH CY DE DK ES FI FR GB GR IE IT LU MC NL PT SE TR
(OA) BF BJ CF CG CI CM GA GN GW ML MR NE SN TD TG
(AP) GH GM KE LS MW MZ SD SL SZ TZ UG ZW
(EA) AM AZ BY KG KZ MD RU TJ TM
Main International Patent Class: G06F-009/46
International Patent Class: G06F-017/30; G06F-017/60
Publication Language: English
Filing Language: English
Fulltext Availability:
Detailed Description
Claims
Fulltext Word Count: 84642

English Abstract
A comprehensive electronic business support system comprises three layers: (1) the business layer, including various smart components which unify data and business processes across all customer interactions; (2) the integration layer, including various communications messaging interfaces and enterprise application integration adpaters, which provide a flexible, automated, and process driven solution for integrating across business applications and operations support systems; and (3) the presentation layer, including various customer views, which are presented via particular business portals. A smart component server provides the core services and comprehensive business process logic required to successfully conduct business online. The communications messaging interfaces integrate with back-office systems for functions such as billing, provisioning, and interconnection.

French Abstract
Systeme de support global pour commerce electronique qui comporte trois couches, dont (1) la couche commerciale, comportant divers composants intelligents qui homogeneisent les processus de donnees et commerciaux pour toutes les interactions avec les clients, (2) la couche d'integration, y compris diverses interfaces de messagerie de communications et adaptateurs d'integration d'applications d'entreprise, qui fournissent une solution souple, automatisee et commandee par les processus pour l'integration des applications inter-commerciales et des systemes de support d'operations, et (3) la couche de presentation, y compris diverses presentations a l'intention des utilisateurs, qui sont presentees via des portails commerciaux particuliers. Un serveur a composants intelligents fournit les services cles et une logique globale de processus commerciaux, requis pour effectuer avec succes du commerce en ligne. Les interfaces de messagerie de communications s'integrent dans des systemes d'arriere-guichet pour des fonctions telles que la

facturation, l'approvisionnement et l'interconnexion.

Legal Status (Type, Date, Text)
Publication 20011011 A2 Without international search report and to be republished upon receipt of that report.
Correction 20011227 Corrections of entry in Section 1: under (72, 75) replace "POKOTYLO, Vadim [-/US]" by "POKOTYLO, Vadim [MD/US]" and "BAKHRU, Girish [-/US]" by "BAKHRU, Girish [IN/US]"
Republication 20011227 A2 Without international search report and to be republished upon receipt of that report.
Correction 20011227 Corrections of entry in Section 1:
Examination 20020214 Request for preliminary examination prior to end of 19th month from priority date
Search Rpt 20020718 Late publication of international search report
Republication 20020718 A3 With international search report.
Republication 20020718 A3 Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

Fulltext Availability:
Detailed Description

Detailed Description
... 5 7
SECTION 2. DATA STRUCTURE
This section describes the data structure for the eBusiness support system which includes various database tables.

2. 1. AREA **TABLE** DESCRIPTIONS
All **tables** of the database are grouped according to their area. Each **table** lists the name of each field attribute, whether or not the field is required, the type and length, and a description for each field.

(1) Populating **tables** manually
Every row in most database **tables** uses an Object Identifier (OID) as the
primary key. To manually populate a row in a database, the entry for the row's
OID must be determined for that area, using the SEQ **table** . The SEQ **table**
contains an entry for each area of the application (the SEQ- **NAME attribute**)
and a **VALUE** . The value is the next number to be used as an OID for that area, and is the number to be used as the OID...

...once this number is used, it must also be manually incremented, so that the next entry made will use a unique number.

(2) Type
Some **tables** have a TYPE field, a class indicator field used by the persistence layer to determine the subclass for a particular object. If the database is the correct value, the Topl-ink Builder console is used to view the inheritance properties for the superclass object.

(3) Write Lock
Most **tables** have a WRITE-LOCK field. This field works with the persistence layer to provide an optimistic lock that prevents access to a field if it is in the process of updating.

(4) Primary Keys
The field name OID denotes the primary key for a **table** . Any other field that uses "-OID" in its name is a foreign key.

5 9
(5) Dates.
All unset dates are treated as Y.

AGENT...Name Allows Type Description
Nulls?
6 9
Table 2. 20 represents an invoice charge item that represents usage charges imported from an external billing system.

Table 2 INVOICE-USAGE-CHARGE-ITEM
Attribute Name Allows Type Description
Nulls?
OID N NUMBER (18) Object identifier for the invoice
usage charge
START DT N...

...domain type for the "to" service
identifier
FROM-SVC- DOMAIN-CD N NUMBER (9) 'Code indicating the service
domain type for the "from"
ervice identifier
Table 2. 21. is a collection of attributes representing Adjustment
Request to a Customer Bill.
Table 2 BILL-ADJMNT-RQST
Attribute Name Allows Type Description
Nulls?
OID IN INUMBER (18) JObject identifier for the bill adjustment
70
request
BLNG -POINT OID...

11/5,K/38 (Item 29 from file: 349)
DIALOG(R)File 349:PCT FULLTEXT
(c) 2003 WIPO/Univentio. All rts. reserv.

00835791 **Image available**
SYSTEM AND METHOD FOR CREATING A SEMANTIC WEB AND ITS APPLICATIONS IN BROWSING, SEARCHING, PROFILING, PERSONALIZATION AND ADVERTISING
SYSTEME ET PROCEDE DE CREATION D'UN WEB SEMANTIQUE ET DE SES APPLICATIONS DANS L'EXPLORATION, LA RECHERCHE, LE PROFILAGE, LA PERSONNALISATION ET LA PUBLICITE
Patent Applicant/Assignee:
TAALEE INC, Town Place, 263 W. Clayton Street, Suite 5, Athens, GA 30601, US, US (Residence), US (Nationality)
Inventor(s):
SHETH Amit, 1140 Laurel Pointe, Bogart, GA 30622, US,
AVANT David, 1340 Malcolm Bridge Road, Bogart, GA 30622, US,
BERTRAM Clemens, 1287 Cedar Shoals Drive, Apartment 820, Athens, GA 30605, US,
Legal Representative:
FLOAM D Andrew (et al) (agent), Needle & Rosenberg, P.C., The Candler Building, Suite 1200, 127 Peachtree Street, Atlanta, GA 30303-1811, US,

Patent and Priority Information (Country, Number, Date):
Patent: WO 200169428 A1 20010920 (WO 0169428)
Application: WO 2001US8188 20010314 (PCT/WO US0108188)
Priority Application: US 2000189528 20000315; US 2000645301 20000824
Designated States: AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CO CR CU CZ DE DK DM DZ EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR LS LT LU LV MA MD MG MK MN MW MX MZ NO NZ PL PT RO RU SD SE SG SI SK SL TJ TM TR TT TZ UA UG UZ VN YU ZA ZW
(EP) AT BE CH CY DE DK ES FI FR GB GR IE IT LU MC NL PT SE TR
(OA) BF BJ CF CG CI CM GA GN GW ML MR NE SN TD TG
(AP) GH GM KE LS MW MZ SD SL SZ TZ UG ZW
(EA) AM AZ BY KG KZ MD RU TJ TM
Main International Patent Class: G06F-017/00
Publication Language: English
Filing Language: English
Fulltext Availability:
Detailed Description

Claims
Fulltext Word Count: 10406

English Abstract
A system and method for creating a database of metadata (metabase) [14] of a variety of digital media content, including TV and radio content delivered on Internet. This semantic-based method captures and enhances domain or subject specific metadata of digital media content, including the specific meaning and intended use of original content, including the specific meaning and intended use of original content. To support semantics, a WorldModel [11] is provided that includes specific domain knowledge, ontologies as well as a set of rules relevant to the original content. The metabase [14] may also be dynamic in that it may track changes to the any variety of accessible content, including live and archival TV and radio programming.

French Abstract
L'invention concerne un systeme et un procede de creation d'une base de donnees de meta-donnees (meta-base) [14] d'une variete de contenus mediatiques numeriques, y compris le contenu televisuel et radiophonique delivre sur Internet. Ce procede fonde sur la semantique capture et ameliore des meta-donnees de sujets ou de domaines specifiques de contenus mediatiques numeriques, y compris la signification specifique et l'utilisation projetee de contenus originaux. Afin de supporter la semantique, un modele du monde [11] comprend des connaissances de domaines specifiques, des ontologies, ainsi qu'un ensemble de regles pertinent pour le contenu original. La meta-base [14] peut egalement etre dynamique en ce qu'elle peut pister des changements dans n'importe quel type de contenu accessible, y compris la programmation televisuelle et radiophonique en direct et d'archives.

Legal Status (Type, Date, Text)
Publication 20010920 A1 With international search report.
Examination 20011213 Request for preliminary examination prior to end of 19th month from priority date

Fulltext Availability:
Detailed Description

Detailed Description
... Web sites and retrieve digital media metadata from selected pages.

1 5 An extractor proararn takes HTML pages and extraction rules as input and generates **XML** assets such as that shown in Fig. 6. These generated assets contain **values** for each **attribute** **name** belonging to the domain of that Web site. Once created, the assets are sent to a Metabase Agent that is in charge of enhancing and...

...order to enhance the assets, the Metabase Agent uses information stored in the WorldModcl as well as a Knowledgebase. The Knowledoebase is a collection of **tables** containing domain-specific information and relationships. After insertion into the rnetabase, the assets are then ready to be searched.

The purpose of a WebCrawler is...

11/5,K/39 (Item 30 from file: 349)
DIALOG(R)File 349:PCT FULLTEXT

00828886
RDL SEARCH ENGINE
MOTEUR DE RECHERCHE RDL
Patent Applicant/Assignee:
E-NUMERATE SOLUTIONS INC, 8201 Greensboro Drive, McLean, VA 22102, US, US (Residence), US (Nationality)
Inventor(s):

DAVIS Russell T, Bethesda, MD, US,
Legal Representative:
GARRETT Arthur S (et al) (agent), Finnegan, Henderson, Farabow, Garrett & Dunner, L.L.P., 1300 I Street, N.W., Washington, DC 20005-3315, US,
Patent and Priority Information (Country, Number, Date):
Patent: WO 200161568 A2 20010823 (WO 0161568)
Application: WO 2001US5268 20010216 (PCT/WO US0105268)
Priority Application: US 2000183152 20000217
Designated States: AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CR CU CZ DE DK DM DZ EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR LS LT LU LV MA MD MG MK MN MW MX MZ NO NZ PL PT RO RU SD SE SG SI SK SL TJ TM TR TT TZ UA UG UZ VN YU ZA ZW
(EP) AT BE CH CY DE DK ES FI FR GB GR IE IT LU MC NL PT SE TR
(OA) BF BJ CF CG CI CM GA GN GW ML MR NE SN TD TG
(AP) GH GM KE LS MW MZ SD SL SZ TZ UG ZW
(EA) AM AZ BY KG KZ MD RU TJ TM
Main International Patent Class: G06F-017/30
Publication Language: English
Filing Language: English
Fulltext Availability:
Detailed Description
Claims
Fulltext Word Count: 15802

English Abstract
Methods and systems consistent with the present invention provide a means for searching numerical data across networks such as the Internet, and removing the middle layer of query engines or servers used by conventional systems in retrieving data from relational databases over the Internet. The methods and systems in accordance with the present invention also provide a means for tying millions of computers together into a single database, thereby a query introduced to the system returns a table of data as a single database is capable of providing. Furthermore, the methods and systems consistent with the present invention provide the means for performing navigational, line item (or record-level), semantic, numerical, transformational, arithmetic, time-dependent, and cost based queries on numerical data. In addition, a user may also conduct select queries between unrelated databases.

French Abstract
La presente invention concerne des procedes et des systemes fournissant un moyen pour rechercher des donnees numeriques dans des reseaux tels qu'Internet, et eliminer la couche centrale des moteurs de requete ou des serveurs utilises par les systemes classiques pour extraire des donnees a partir de bases de donnees relationnelles dans Internet. Ces procedes et ces systemes fournissent egalement un moyen pour relier des millions d'ordinateurs ensemble dans une seule base de donnees, de maniere a ce qu'une requete introduite dans le systeme restitue une table de donnees comme une seule base de donnees. Par ailleurs, ces procedes et ces systemes fournissent un moyen pour realiser des requetes de navigation, d'article de ligne (ou de niveau d'enregistrement), semantiques, numeriques, transformationnelles, d'arithmetique, dependantes du temps, et des requetes basees sur le cout sur des donnees numeriques. De plus, l'utilisateur peut egalement gerer des requetes de selection entre les bases de donnees non reliees.

Legal Status (Type, Date, Text)
Publication 20010823 A2 Without international search report and to be republished upon receipt of that report.
Examination 20011129 Request for preliminary examination prior to end of 19th month from priority date
Correction 20020411 Corrected version of Pamphlet: pages 1-43, description, replaced by new pages 1-39; pages 44-62, claims, replaced by new pages 40-56; pages 1/15-15/15, drawings, replaced by new pages 1/15-15/15; due to late transmittal by the receiving Office
Republication 20020411 A2 Without international search report and to be

repu[redacted]ished upon receipt of that repo[redacted].

Fulltext Availability:
Detailed Description

Detailed Description
... be copied to a local source and the restriction on cataloging numbers is relaxed.

At this step, the index of documents is created. A standard **XML** processor is used to proceed through each document, element by element, and then attribute by attribute to create an index of the major elements and attributes (step 1108).

In **XML** , "element" text values are the values appearing between opening and closing tags (e.g., "< data-source > US Census Bureau < data-source>" would have an element...

...tag itself For example, "< lineitem> li-legend = Total'Revenues'> < line-item >" would have an attribute named "lijegend" which has a value of "Total Revenues."
The **XML** processor collects both elements and **attributes** into **name** / **value pairs** and calls a software routine (or method) as each is completed. The "handler" method takes the URL, element/attribute name, and value and creates a new record in the index data **table** with those three values. The "cache" RDL documents are also handed off to an additional handler which collects all the attributes and elements into a relational **table** (step I I I 1). During the search process, the data query processor 204 will use this relational cache to create the RDL elements in...of items that may be indexed, for example, element names (what is found within the "< > ... s), element values (what is found between the 64< >995s)@ **attribute names** (what is found within the "< >`s and before an "="), **attribute values** (what is found between the "< > ... s and after "="), and various types of metadata outside of the documents (e.g., number of hits, response time of...

...10 shows an index with four (4) columns, other implementations may have different numbers of colunms. Discussed below are three different implementations of an index **table** , each with a different set of tag type infori-nation. Note that **attribute names** and legend **names** are mixed together in the same column: there is no overlap between **attribute** and element **names** in RDL. One skilled in the art with **XML** experience will recognize that this
30
efficiency can be relaxed to accommodate languages without this feature by adding an additional
column called "tag type."
In one implementation, **attribute names** and **values** are recorded. This leads to the smallest index: easiest to create and fastest to search. To really test a query, however, requires going to the...

11/5,K/40 (Item 31 from file: 349)
DIALOG(R)File 349:PCT FULLTEXT

00824139 **Image available**
METHOD AND SYSTEM FOR TESTING INTERNET-BASED APPLICATIONS
PROCEDE ET SYSTEME DE MISE A L'ESSAI D'APPLICATIONS BASEES SUR INTERNET
Patent Applicant/Assignee:
MOBILEQ CANADA INC, 175 Bloor Street East, 12th Floor, Toronto, Ontario M4W 3R8, CA, CA (Residence), CA (Nationality), (For all designated states except: US)
Patent Applicant/Inventor:
CANARAN Vishvas, 25 The Esplanade, Suite 3015, Toronto, Ontario M5E 1W4, CA, CA (Residence), CA (Nationality), (Designated only for: US)
PERLA Jesse, 38 Elm Street, Suite 3108, Toronto, Ontario M5G 2K5, CA, CA (Residence), CA (Nationality), (Designated only for: US)

Legal Representative:
PILLAY Kevin (agent), Fasken Martineau DuMoulin LLP, Toronto Dominion Bank Tower, Suite 4200, P.O. Box 20, Toronto-Dominion Centre, Toronto, Ontario M5K 1N6, CA,

Patent and Priority Information (Country, Number, Date):
Patent: WO 200157671 A1 20010809 (WO 0157671)
Application: WO 2001CA147 20010131 (PCT/WO CA0100147)
Priority Application: CA 2297597 20000131; CA 2297711 20000131; CA 2297596 20000131
Designated States: AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CR CU CZ DE DK DM DZ EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR LS LT LU LV MA MD MG MK MN MW MX MZ NO NZ PL PT RO RU SD SE SG SI SK SL TJ TM TR TT TZ UA UG US UZ VN YU ZA ZW
(EP) AT BE CH CY DE DK ES FI FR GB GR IE IT LU MC NL PT SE TR
(OA) BF BJ CF CG CI CM GA GN GW ML MR NE SN TD TG
(AP) GH GM KE LS MW MZ SD SL SZ TZ UG ZW
(EA) AM AZ BY KG KZ MD RU TJ TM
Main International Patent Class: G06F-011/36
Publication Language: English
Filing Language: English
Fulltext Availability:
Detailed Description
Claims
Fulltext Word Count: 9157

English Abstract
A method for testing a web-based application comprising a plurality of forms such method comprising the steps of defining a form flow of ones of a plurality of forms, defining test parameters for a test; creating a test script file defined by a test parameters and a form flow, and, generating a plurality of sets of form requests in accordance with a test script, wherein ones of a plurality of sets of form requests are generated for each permutation of test parameters.

French Abstract
L'invention concerne un procede permettant de mettre a l'essai une application basee sur le Web qui englobe plusieurs formes. Ce procede consiste a definir un flux de forme parmi certaines formes d'une pluralite de formes, definir des parametres d'essai pour un essai, creer un fichier de scenario d'essai defini par des parametres d'essai et ledit flux de forme, et generer plusieurs series de demandes de formes en fonction du scenario d'essai, une desdites series de demandes de formes etant generee pour chaque permutation desdits parametres d'essai.

Legal Status (Type, Date, Text)
Publication 20010809 A1 With international search report.
Publication 20010809 A1 Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.
Examination 20011101 Request for preliminary examination prior to end of 19th month from priority date

Fulltext Availability:
Claims

Claim
... access 5 the data in a distributed manner across the Internet. External hosting requires remote execution of data. The implementation of Unicode as part of **Extensible Markup** Language (**XML**) and Extensible Stylesheet Language (XSQ standard is one of the reasons allowing applications to be made international. Unicode is a standard for interchanging, processing, and...

...method for testing the application comprising the steps of(a) creating a test script file defined by test cases in a test script using an **XML** based markup language; and
(b) providing a test script executive for running said test script file.
hi accordance with a further aspect of the invention...XSL stylesheet

interpreter with the appropriate HTML and the stylesheet to create the formatted markup of the appropriate MIME type. In some cases both the **XML** input and the XSL stylesheet may be provided to the client browser to interpret if the client has an XSL built. By way of background, **Extensible Markup** Language, abbreviated **XML** , describes a class of data objects called dt-xi-nl-doe **XML** documents and partially describes the behavior of computer programs which process them. **XML** is an application profile or
restricted form of SGML, the Standard Generalized Markup Language. By construction, **XML** documents are conforming SGML documents. **XML** documents are made up of storage units called entities, which contain either 1 5 parsed or unparsed data. Parsed data is made up of characters ...

...which form character data dt-chardata, and some of which form markup. Markup encodes a description of the document's storage layout and logical structure. **XML** provides a mechanism to impose constraints on the storage layout and logical structure. A software module called an **XML** processor is used to read **XML** documents and provide access to their content and structure. It is assumed that an **XML** processor is doing its work on behalf of another module, called the application, which resides on the web server 108. Each **XML** document has both a logical and a physical structure. Physically, the document is composed of the units called entities. An entity may refer to other...

...all of which are indicated in the document by explicit markup. The logical and physical structures must nest properly.
6
SUBSTITUTE SHEET (RULE 26)
Each **XML** Document contains one or more elements, the boundaries of which are delimited by start-tags and end-tags. Each element has a type, identified by name, sometimes called its "generic identifier" (GI), and may have a set of attribute specifications. Each **attribute** specification has a **name** and a value. Style Sheets can be associated with an **XML** document by using a processing instruction whose target is **xml** -stylesheet. This processing instruction follows the behavior of HTML 4 The **xml** -stylesheet processing instruction is parsed in the same way as a start-tag, with the exception that entities other than predefined entities I 0 must...

...stylesheet to be applicable to a wide class of documents that have similar source tree structures. Each stylesheet describes rules for presenting a class of **XML** source documents. An XSL stylesheet processor accepts a document or data in **XML** and an XSL stylesheet and produces the presentation of that **XML** source content as intended by the stylesheet. There are two sub-processes to this presentation process First, a result tree is constructed from the **XML** source tree. Second, the result tree is interpreted to produce a formatted presentation on a display, on paper, in speech or onto other media. The...

...obtain
8
SUBSTITUTE SHEET (RULE 26)
data for the URL from the databases 218 and 216. The data server 206 retrieves the appropriate data in **XML** , and forwards it to the run-time 204. The run-time 204 adds runtime and directory information to the **XML** data. The data structure that is built or populated by the run-time 204 from HML is termed RML, the structure of which will be...

...shown at numeral 300 a defined schema or data structure for the elements contained in an HML application. All of the elements are described as **XML** -based schemas including attributes and elements. The first element of the HML 300 is an Application Element 301 which is a root element comprising the...contained within an events element which contains multiple
component elements 309;
Multiple Directory elements 308 containing information to connect to directory type data. Contain a **name** **attribute** . Connection Element 303

None.
<batch> element
Represents a list of scripts to be performed in batch. This element is optional and repeatable.
Parent Elements
Element Remarks
<application>
Attributes
Attribute Description
Name Item's distinguished name. Description Description of the batch.
1 0 Text Content
None.
Child Elements
Order lement Remarks

1 <batchmember> Optional, Repeatable.
Remarks
The...

11/5,K/41 (Item 32 from file: 349)
DIALOG(R)File 349:PCT FULLTEXT

00809291 **Image available**
SYSTEM AND METHOD FOR THE STORAGE, INDEXING AND RETRIEVAL OF XML DOCUMENTS USING RELATIONAL DATABASES
SYSTEME ET PROCEDE DESTINES A STOCKER, INDEXER ET EXTRAIRE DES DOCUMENTS XML AU MOYEN DE BASES DE DONNEES RELATIONNELLES
Patent Applicant/Assignee:
B-BOP ASSOCIATES INC, Suite 100, One Bay Plaza, 1350 Old Bayshore Highway, Burlingame, CA 94010, US, US (Residence), US (Nationality)
Inventor(s):
DODDS David, 16 Old Barn Road, Stamford, CT 06905, US,
KUO Larry, 120 Morning Star Drive, San Jose, CA 95131, US,
SENGUPTA Soumitra, 15 First Street, Apt. 5, Stamford, CT 06905, US,
LINDSEY Bill, 2203 Hastings Drive, Apt. 28, Belmont, CA 94002, US,
Legal Representative:
LOHSE Timothy W (agent), Gray Cary Ware & Freidenrich LLP, 400 Hamilton Avenue, Palo Alto, CA 94301-1825, US,
Patent and Priority Information (Country, Number, Date):
Patent: WO 200142881 A2-A3 20010614 (WO 0142881)
Application: WO 2000US42665 20001206 (PCT/WO US0042665)
Priority Application: US 99169101 19991206
Designated States: AE AL AM AT AU AZ BA BB BG BR BY CA CH CN CR CU CZ DE DK DM EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR LS LT LU LV MA MD MG MK MN MW MX NO NZ PL PT RO RU SD SE SG SI SK SL TJ TM TR TT TZ UA UG UZ VN YU ZA ZW
(EP) AT BE CH CY DE DK ES FI FR GB GR IE IT LU MC NL PT SE TR
(OA) BF BJ CF CG CI CM GA GN GW ML MR NE SN TD TG
(AP) GH GM KE LS MW MZ SD SL SZ TZ UG ZW
(EA) AM AZ BY KG KZ MD RU TJ TM
Main International Patent Class: G06F-017/30
Publication Language: English
Filing Language: English
Fulltext Availability:
Detailed Description
Claims
Fulltext Word Count: 6031

English Abstract
A system (34) and method for assigning attributes to XML documents to facilitate their storage and retrieval in relational databases (54). A converter (50) accepts XML documents, processes them and outputs relational data about the XML documents which is stored in the relational database (54). A searcher (52) using SQL query engines performs queries to retrieve the documents and sends the results of the query in XML form to the users.

French Abstract

L'invention concerne un systeme et un procede permettant d'attribuer des attributs a des noeuds de document XML pour faciliter leur stockage dans des bases de donnees relationnelles et l'extraction et la reconstruction consecutives de noeuds et de fragments pertinents dans l'ordre original du document. Etant donne que ces requetes sont effectuees a l'aide de modules de requete de la base de donnees relationnelle, leur vitesse d'execution est significativement plus rapide que lorsque d'autres systemes sont utilises tels que des bases de donnees orientees objet. En outre, ce procede peut etre porte sur toutes les plates-formes de vendeurs et peut donc etre mise en place sur des sites client sans qu'il soit necessaire de proceder a des investissements supplementaires dans un logiciel de base de donnees.

Legal Status (Type, Date, Text)
Publication 20010614 A2 Without international search report and to be republished upon receipt of that report.
Examination 20011018 Request for preliminary examination prior to end of 19th month from priority date
Search Rpt 20020110 Late publication of international search report
Republication 20020110 A3 With international search report.

Fulltext Availability:
Detailed Description
Claims

Detailed Description

... simple SQL queries to retrieve the content desired by the user.

In accordance with another aspect of the invention, a computer system for storing an **XML** document using a relational database is provided wherein the system comprises a converter that receives an **XML** document and generates relational database **tables** based on the structure of the **XML** document. The converter further comprises a software module that generates a unique **name attribute** for each node in the **XML** document, a software module that generates a path attribute for a particular node of the **XML** document wherein the path attribute comprises a list of the **name attributes** for the one or more nodes from the particular node to a root node of the **XML** document, a software module that generates an order attribute for the particular node, the order attribute comprising an enumerated order of the particular node from the root node to the particular node, and a software module that generates a NodeValue **attribute** containing a **value** of the particular node. Collectively these attributes are called encodings that result in efficient storage, indexing and searching of **XML** documents without destroying the underlying hierarchical structure of the documents.

In accordance with yet another aspect of the invention, a data structure that stores a...

Claim

... an XML document and generates a relational database table based on the XML document;
the converter further comprising a software module that generates a unique **name attribute** for each node in the **XML** document, a software module that generates a path attribute for a particular node of the **XML** document wherein the path attribute comprises a list of the **name attributes** for the one or more nodes from the particular node to a root node of the **XML** document, a software module that generates an order attribute for the particular node, the order attribute comprising an enumerated order of I I the particular node from the root node to the particular node, and a software module that generates a NodeValue **attribute** containing a **value** of the particular node.

13 A method for manipulating an XML document using a relational database, comprising:
generating a relational database table based on an...

11/5,K/42 (Item 33 from file: 349)
DIALOG(R)File 349:PCT FULLTEXT

00795116 **Image available**

A METHOD AND SYSTEM FOR AUTOMATICALLY STRUCTURING CONTENT FROM UNIVERSAL MARKED-UP DOCUMENTS

PROCEDE ET SYSTEME DE STRUCTURATION AUTOMATIQUE DE CONTENU A PARTIR DE DOCUMENTS UNIVERSELS MARQUES

Patent Applicant/Assignee:
THE SHOPPER INC, 1209 Orange St., Wilmington, DE 19801, US, US (Residence), US (Nationality), (For all designated states except: US)
Patent Applicant/Inventor:
FISHMAN Alon, Hadganim St. 19, 53482 Givatayim, IL, IL (Residence), IL (Nationality), (Designated only for: US)
ENOSHI Ari, Anilevitch St. 58, 51489 Bnei Brak, IL, IL (Residence), IL (Nationality), (Designated only for: US)
RAN Udi, Gezer St. 6, 64376 Tel Aviv, IL, IL (Residence), IL (Nationality), (Designated only for: US)
Legal Representative:
FRIEDMAN Mark M (agent), Beit Samueloff, Haomanim St. 7, 67897 Tel Aviv, IL,

Patent and Priority Information (Country, Number, Date):
Patent: WO 200127712 A2-A3 20010419 (WO 0127712)
Application: WO 2000IL648 20001012 (PCT/WO IL0000648)
Priority Application: US 99158854 19991012
Designated States: AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CR CU CZ DE DK DM DZ EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR LS LT LU LV MA MD MG MK MN MW MX MZ NO NZ PL PT RO RU SD SE SG SI SK SL TJ TM TR TT TZ UA UG US UZ VN YU ZA ZW
(EP) AT BE CH CY DE DK ES FI FR GB GR IE IT LU MC NL PT SE
(OA) BF BJ CF CG CI CM GA GN GW ML MR NE SN TD TG
(AP) GH GM KE LS MW MZ SD SL SZ TZ UG ZW
(EA) AM AZ BY KG KZ MD RU TJ TM
Main International Patent Class: G06F-017/00
International Patent Class: G06F-007/00
Publication Language: English
Filing Language: English
Fulltext Availability:
Detailed Description
Claims
Fulltext Word Count: 19753

English Abstract
A system and method for automatic, knowledge-based processing and structuring of information from marked-up documents such as SGML, XML, HTML based documents. The system includes three parts: an offer processing (1000), an offer presentation (2000) and a database system (3000) that the offer processing (1000) and the offer presentation (2000) use. The offer processing (1000) builds anoffer database in the database system (3000) by accessing data sources such as Internet sites, intranet files, retrieving pages and processing them. The offer presentation (2000) allows users to access the offer database and retrieve information from it. The offer presentation (2000) also includes servers that may be accessed by different web-enabled means such as web browers, cellular phone, PDA, digital TV, Internet appliances, voice acrivated user interfaces.

French Abstract
L'invention concerne un systeme et un procede de traitement et de structuration automatiques et bases sur la connaissance d'informations de documents universels marques (tout type de document qui contient des informations structurelles, de presentation et semantiques (telles que des balises) associees a un contenu, p. ex. des documents SGML / XML / HTML). La presente invention peut s'appliquer a Internet ou a Intranet, Extranet et a d'autres sources informations en reseau. La technique

fondamentale de la presente invention est un systeme et un procede de traitement d'offres. Le systeme de traitement d'offres rassemble, stocke, traite, extrait et presente des offres d'informations, de produits ou de services. Le systeme rassemble automatiquement des offres d'un tres grand nombre de sites d'informations, et effectue cette operation sans recours a des reglages specifiques au site ou a des connaissances relatives a l'organisation specifique des donnees d'offres dans des pages d'un site donne. Une caracteristique optionnelle de la presente invention est un systeme frontal destine a presenter des offres structurees a des utilisateurs ou a des tiers. Une application du procede et systeme de traitement d'offres peut par exemple etre un systeme et procede d'achats comparatifs. Cette application est un moteur d'achats comparatifs de tres grande echelle et automatique qui rassemble des offres de produits ou de services provenant d'un nombre indefini de vendeurs par Internet ou de fournisseurs de services, aussi bien sur le plan mondial (par le Web) que local (clic & mortier). Les offres d'emploi, immobilieres, de logement, d'entreprises a entreprises (B2B) etc. sont autant d'exemples d'informations qui peuvent etre rassemblees.

Legal Status (Type, Date, Text)
Publication 20010419 A2 Without international search report and to be republished upon receipt of that report.
Search Rpt 20011018 Late publication of international search report
Republication 20011018 A3 With international search report.
Fulltext Availability:
Claims

Claim
... the web
page's data structure, apart of the trivial assumption that the page is being
presented in the generally used HTML format (later an **XML** format option will be added).
C Algorithm's output
If the algorithm succeeds to fetch one or more product offerings in the relevant product category...almost all web pages today) contains no definitive meaningful data structure. This was one of the main reasons for the development of structured formats like **XML** . When looking on the general context this might be true. This novel algorithm, however, finds structure in the structure less web pages sub-population of...

...we should take some practical assumptions that are not a big constrain in the consumer e-commerce environment.
Those assumptions are:
A. 1 1 HTML/ **XML**
Merchant content may be presented as an f ITM L (in the future also **XML**) document (applicable to almost all the textual content on the WWW).
A. 1. 2. 11 Weak order of the content layou
We can't take...represents the certainty level of this assumption (it is like probability
value, but we call it Likelihood-mark because it is not a real probability **value**).
A 5 **Attribute** set
A set S of **attributes** .
A 6 Single **value** field
A field F will be marked as "single value field" if in a legitimate product offer record. there will be just one value associated with that field.
A 7 Combination value field
A field F will be marked as "Combination **value** **attribute** " if in a legitimate product offer record, this field can get a list of different values, and the meaning of this list is that the...TA such that:
* S' is a legitimate product record in the product category PC. + PA is a price attribute (that means that PA is an **attribute** , PA. Field - **name** @ "Price").

```
File   8:Ei Compendex(R) 1970-2003/Nov W2
         (c) 2003 Elsevier Eng.  Info. Inc.
File  35:Dissertation Abs Online 1861-2003/Oct
         (c) 2003 ProQuest Info&Learning
File 202:Info. Sci. & Tech. Abs. 1966-2003/Nov 17
         (c) 2003 EBSCO Publishing
File  65:Inside Conferences 1993-2003/Nov W3
         (c) 2003 BLDSC all rts. reserv.
File   2:INSPEC 1969-2003/Nov W2
         (c) 2003 Institution of Electrical Engineers
File 233:Internet & Personal Comp. Abs. 1981-2003/Jul
         (c) 2003, EBSCO Pub.
File  94:JICST-EPlus 1985-2003/Nov W3
         (c)2003 Japan Science and Tech Corp(JST)
File 603:Newspaper Abstracts 1984-1988
         (c)2001 ProQuest Info&Learning
File 483:Newspaper Abs Daily 1986-2003/Nov 17
         (c) 2003 ProQuest Info&Learning
File   6:NTIS 1964-2003/Nov W3
         (c) 2003 NTIS, Intl Cpyrght All Rights Res
File 144:Pascal 1973-2003/Nov W2
         (c) 2003 INIST/CNRS
File 434:SciSearch(R) Cited Ref Sci 1974-1989/Dec
         (c) 1998 Inst for Sci Info
File  34:SciSearch(R) Cited Ref Sci 1990-2003/Nov W2
         (c) 2003 Inst for Sci Info
File  99:Wilson Appl. Sci & Tech Abs 1983-2003/Oct
         (c) 2003 The HW Wilson Co.
File 583:Gale Group Globalbase(TM) 1986-2002/Dec 13
         (c) 2002 The Gale Group
File 266:FEDRIP 2003/Sep
         Comp & dist by NTIS, Intl Copyright All Rights Res
File  95:TEME-Technology & Management 1989-2003/Nov W1
         (c) 2003 FIZ TECHNIK
File 438:Library Lit. & Info. Science 1984-2003/Oct
         (c) 2003 The HW Wilson Co

Set     Items   Description
S1      17297   XML OR (EXTENSIBLE OR XTENSIBLE)()(MARKUP OR MARK()UP)
S2        554   NAME? ?(3N)ATTRIBUTE? ?
S3       8831   VALUE? ?(3N)ATTRIBUTE? ?
S4         19   S2(5N)S3(5N)(PAIR??? OR COUPLE? ? OR SET? ? OR GROUP???? OR
                TOKEN? ?)
S5     634177   TABLE? ? OR TABLESPACE? ? OR LUT? ? OR HASHTABLE? ?
S6        358   CODESPACE? ? OR CODE()SPACE? ?
S7          4   S1 AND S2 AND S3
S8          0   S1 AND S4
S9          0   S1 AND S6
S10         0   S7 AND S5:S6
S11         4   RD S7 (unique items)
```

11/5/1 (Item 1 from file: 2)
DIALOG(R)File 2:INSPEC

6040799 INSPEC Abstract Number: C9811-7240-008
Title: An efficiently updatable index scheme for structured documents
Author(s): Kanemoto, H.; Kato, H.; Kinutani, H.; Yoshikawa, M.
Author Affiliation: Graduate Sch. of Inf. Sci., Nara Inst. of Sci. & Technol., Japan
Conference Title: Proceedings Ninth International Workshop on Database and Expert Systems Applications (Cat. No.98EX130) p.991-6
Editor(s): Tjoa, A.M.; Wagner, R.R.
Publisher: IEEE Comput. Soc, Los Alamitos, CA, USA
Publication Date: 1998 Country of Publication: USA xix+1023 pp.
ISBN: 0 8186 8353 8 Material Identity Number: XX98-02383
U.S. Copyright Clearance Center Code: 0 8186 8353 8/98/$10.00
Conference Title: Proceedings Ninth International Workshop on Database and Expert Systems Applications
Conference Sponsor: IEEE Comput. Soc.; DEXA Assoc.; Austrian Comput. Soc.; Res. Inst. Appl. Knowledge Process (FAW); Univ. Vienna
Conference Date: 26-28 Aug. 1998 Conference Location: Vienna, Austria
Language: English Document Type: Conference Paper (PA)
Treatment: Practical (P)
Abstract: We propose an efficiently updatable index scheme for **XML** documents. This index scheme consists of four types of indices. Content index manages occurrence positions of words, element **names** , **attribute names** and **attribute values** . Local structure index manages logical structure of each document. Global structure index summarizes logical structure of document instances for a DTD. The last one is called structure meta index, and keeps, for each element, the number of local structure indices and global structure indices which include the element. Using our four types of indices, therefore, a wide range of queries over structured documents can be processed efficiently. We compare our indices with existing indices in terms of index size, update cost and retrieval cost. (4 Refs)
Subfile: C
Descriptors: data structures; indexing
Identifiers: updatable index scheme; structured documents; **XML** documents; content index; occurrence positions; element names; **attribute names** ; **attribute values** ; logical structure; structure meta index; index size; update cost; retrieval cost
Class Codes: C7240 (Information analysis and indexing); C6120 (File organisation)

11/5/2 (Item 1 from file: 233)
DIALOG(R)File 233:Internet & Personal Comp. Abs.

00614593 00NC11-104
XHTML: crossroads of HTML and XML **-- Sloppy coding habits just won't work with XHTML. We'll help clean up your language**
Abualsamid, Ahmad
Network Computing , November 13, 2000 , v11 n22 p155-158, 3 Page(s)
ISSN: 1046-4468
Company Name: World Wide Web Consortium
Languages: English
Document Type: Articles, News & Columns
Geographic Location: United States
Presents a primer on Extensible Hypertext Markup Language (XHTML), defined as Hypertext Markup Language (HTML) 4.0 expressed using an **Extensible Markup** Language (**XML**) Document Type Definition (DTD). Says that XHTML paves the path for the modularization of code snippets that are loaded dynamically to handle various XHTML modules. Mentions that the World Wide Web Consortium's (W3C) XHTML 1.0 specification has the advantages of extensibility, interoperability, and portability. Indicates the differences between XHTML and HTML: tags and attributes must be in lowercase; all XH

elements must be closed; the **attribute** replaces the name **attribute** ; **attribute values** must be quoted, and no minimizati allowed; XHTML documents have some mandatory elements; and must conform to **XML** rules. Includes a screen display. (MEM)
Descriptors: HTML; **XML** ; Programming Language; Object-oriented; Application Development; Standards
Identifiers: World Wide Web Consortium

11/5/3 (Item 1 from file: 94)
DIALOG(R)File 94:JICST-EPlus

05279518 JICST ACCESSION NUMBER: 02A0839997 FILE SEGMENT: JICST-E
A Virtual Database Integration Method for Nuclear Power Plant Engineering.
SEKI HIROSHI (1); SANO HIROKI (1); YUCHI HIROYUKI (1)
(1) Hitachi, Ltd.
Denki Gakkai Genshiryoku Kenkyukai Shiryo, 2002, VOL.NE-02,NO.6-10, PAGE.5-10, FIG.4, TBL.1, REF.4
JOURNAL NUMBER: G0330BAB
UNIVERSAL DECIMAL CLASSIFICATION: 621.039.4
LANGUAGE: Japanese COUNTRY OF PUBLICATION: Japan
DOCUMENT TYPE: Conference Proceeding
ARTICLE TYPE: Original paper
MEDIA TYPE: Printed Publication
ABSTRACT: In nuclear power plant engineering, many functional areas are involved in and much engineering interactions take place. This necessitates the transfer of engineering data between databases owned by each engineering section. This means that the confirmation of the data consistency between databases is very important. To achieve these functions, we developed an information sharing method for the data consistency between databases using **XML** documents. **XML** Schemas are used for the generation of **XML** documents. The **XML** Schemas are automatically generated from the **attribute values** matching results, and map **attribute names** between databases. We found that this **XML** Schema generation method was effective to share the information for the data consistency between databases. (author abst.)
DESCRIPTORS: nuclear power generation; database; plant design; attribute; computer system(architecture)
BROADER DESCRIPTORS: power generation; electric power energy operation; design; property; method
CLASSIFICATION CODE(S): MD01000C

11/5/4 (Item 2 from file: 94)
DIALOG(R)File 94:JICST-EPlus

05002258 JICST ACCESSION NUMBER: 01A0851433 FILE SEGMENT: JICST-E
Synthesis of Heterogeneous Databases using Intelligent Agents.
ABE KAZUHIRO (1); FUKUSHIMA SHIGEKI (1); ONO TAKAHISA (2)
(1) Mitsubishi Electr. Corp.; (2) Tokyo Electr. Power Co.
Jinko Chino Gakkai Zenkoku Taikai Ronbunshu(Proceedings of the Annual Conference of JSAI), 2001, VOL.15th,NO.Vol.1, PAGE.1F1.04,1-2, FIG.5, REF.4
JOURNAL NUMBER: X0580AAA
UNIVERSAL DECIMAL CLASSIFICATION: 681.3:061.68 681.3:007.51 002.5:005
LANGUAGE: Japanese COUNTRY OF PUBLICATION: Japan
DOCUMENT TYPE: Conference Proceeding
ARTICLE TYPE: Short Communication
MEDIA TYPE: Printed Publication
ABSTRACT: In this paper we describe an information sharing system to synthesize heterogeneous databases using intelligent agents, for supporting power system planning. This system has the following features: 1) Operators can make inquiries using natural language. 2) Fuzzy searches by automatic handling of differences in **attribute names** and **values** between databases are supported. 3) Operators have uniform access to information sources in different forms such as RDB

searching, XML -data reference, and inquires to the data source manager. 4) Functions as those for acquiring, accumulating, and utilizing application know-how, such as inquiries for ensuring the desired results, are supported. (author abst.)

DESCRIPTORS: relational data base; agent; information retrieval; electric power equipment; facility planning; application oriented language; attribute; ambiguity; information source

BROADER DESCRIPTORS: database; retrieval; plan; programming language; formal language; language; property

CLASSIFICATION CODE(S): JD03030U; JE08000Z; AC06020S

```
File 275:Gale Group Computer DB(TM) 1983-2003/Nov 17
         (c) 2003 The Gale Group
File 621:Gale Group New Prod.Annou.(R) 1985-2003/Nov 18
         (c) 2003 The Gale Group
File 636:Gale Group Newsletter DB(TM) 1987-2003/Nov 17
         (c) 2003 The Gale Group
File  16:Gale Group PROMT(R) 1990-2003/Nov 17
         (c) 2003 The Gale Group
File 160:Gale Group PROMT(R) 1972-1989
         (c) 1999 The Gale Group
File 148:Gale Group Trade & Industry DB 1976-2003/Nov 18
         (c)2003 The Gale Group
File 624:McGraw-Hill Publications 1985-2003/Nov 17
         (c) 2003 McGraw-Hill Co. Inc
File  15:ABI/Inform(R) 1971-2003/Nov 17
         (c) 2003 ProQuest Info&Learning
File 647:CMP  Computer Fulltext 1988-2003/Nov W2
         (c) 2003 CMP Media, LLC
File 674:Computer News Fulltext 1989-2003/Nov W2
         (c) 2003 IDG Communications
File 696:DIALOG Telecom. Newsletters 1995-2003/Nov 17
         (c) 2003 The Dialog Corp.
File 369:New Scientist 1994-2003/Nov W2
         (c) 2003 Reed Business Information Ltd.
File 810:Business Wire 1986-1999/Feb 28
         (c) 1999 Business Wire
File 813:PR Newswire 1987-1999/Apr 30
         (c) 1999 PR Newswire Association Inc
File 610:Business Wire 1999-2003/Nov 18
         (c) 2003 Business Wire.
File 613:PR Newswire 1999-2003/Nov 18
         (c) 2003 PR Newswire Association Inc

Set     Items   Description
S1     152004   XML OR (EXTENSIBLE OR XTENSIBLE)()(MARKUP OR MARK()UP)
S2       2919   NAME? ?(3N)ATTRIBUTE? ?
S3       8330   VALUE? ?(3N)ATTRIBUTE? ?
S4         19   S2(5N)S3(5N)(PAIR??? OR COUPLE? ? OR SET? ? OR GROUP???? OR
                TOKEN? ?)
S5    1753715   TABLE? ? OR TABLESPACE? ? OR LUT? ? OR HASHTABLE? ?
S6        701   CODESPACE? ? OR CODE()SPACE? ?
S7          7   S1 AND S4
S8         34   S1 AND S2(S)S3
S9          9   S2(S)S3(S)S5:S6
S10         1   S1 AND S9
S11        34   S7:S8 OR S10
S12        19   RD (unique items)
S13        13   S12 NOT PY=2001:2003
S14        34   S1(S)S2(S)S3 OR S1(100N)S2(100N)S3
S15        19   RD (unique items)
S16        15   S15 NOT PY=2001:2003
```

16/3,K/1 (Item 1 from file: 275)
DIALOG(R)File 275:Gale Group Computer DB(TM)

02463670 SUPPLIER NUMBER: 68018201 (USE FORMAT 7 OR 9 FOR FULL TEXT)
Extensible Markup Language Basics.(Technology Information)
Marco, Lou
Enterprise Systems Journal, 15, 12, 49
Dec, 2000
ISSN: 1053-6566 LANGUAGE: English RECORD TYPE: Fulltext; Abstract
WORD COUNT: 2872 LINE COUNT: 00264

... XML document can be displayed in different ways by using different style sheets, or the same style sheet could govern the display of similarly structured **XML** documents.

The nonproprietary nature of XML, combined with its ease of writing, makes XML an ideal format for data exchange among applications.

A Simple XML...

...some time making her document easy to read with judicious use of tabs, white space and blank lines.

XML processing instructions and tags often use **XML attributes** , which are **name** -value pairs separated by an equals sign; the values must be enclosed in quotes (Another difference between **XML** and HTML is that most HTML values do not require the quotes.) The **XML** declaration requires the use of the version and standalone attributes. For now, know that this **XML** declaration states that the document conforms to **XML** version 1.0, and does not require any other documents for parsing its content.

This document does not contain any display or formatting information. Therefore, one would need a style sheet, and a way of telling the **XML** document to use that style sheet to display the **XML** document. Let's use the simple style sheet below, saved as forfirst.css in the same directory as the **XML** document. Notice that the style sheet refers to the tag <First> used in the document.

```
First {display: block; font-size: 36pt;
font-weight: bold; color="OOFFOO";}
```

The second, italicized processing instruction below associates the stylesheet with the **XML** document:

```
<? xml version="1.0"
standalone="yes"?>
<? xml -stylesheet type="text/css"
href="forfirst.css"?>
<First>
My First XML Document
</First>
```

XML Document Components

XML documents are text consisting of data and markups. The data is what the author encodes; the markups tell the XML parser how this data is ...

16/3,K/2 (Item 2 from file: 275)
DIALOG(R)File 275:Gale Group Computer DB(TM)

02455542 SUPPLIER NUMBER: 66889342 (USE FORMAT 7 OR 9 FOR FULL TEXT)
XHTML: Crossroads of HTML and XML -- Sloppy coding habits just won't work with XHTML. We'll help clean up your language.(Internet/Web/Online Service Information)
Abualsamid, Ahmad
Network Computing, 155
Nov 13, 2000
ISSN: 1046-4468 LANGUAGE: English RECORD TYPE: Fulltext
WORD COUNT: 2080 LINE COUNT: 00161

... form elements. To make sure your code works with today's browsers and with existing scripts, you can use both a name and an ID **attribute** .

- **Attribute values** must be quoted, and no minimization is

allowed. A common practice amo[redacted] HTML coders is to leave the qu[redacted]s out when specifying values for elements...

...in XHTML. Some attributes, such as "checked," could be minimized when using several browsers. This also is not valid. You can't have a dangling **attribute** :

< input id="acheckbox" **name** ="acheckbox" checked /> is incorrect.

< input id="acheckbox" name="acheckbox" checked= "checked" /> is correct.

- XHTML documents have some mandatory elements. You no longer can have documents a minimal XTHML document < /title>

</head >

<body >

</body >

</html >

Rules of **XML**

Another key difference is that XHTML documents must conform to **XML** rules. Here are the pertinent **XML** rules for XHMTL developers (for more information, see www. **xml** .com):

- All **XML** documents are well-formed by definition. A well-formed document adheres to the **XML** structure but does not follow a certain DTD. A document following a certain DTD is called valid. In XHTML, the DTD is that of HTML...

16/3,K/3 (Item 3 from file: 275)

DIALOG(R)File 275:Gale Group Computer DB(TM)

02396193 SUPPLIER NUMBER: 61953375 (USE FORMAT 7 OR 9 FOR FULL TEXT)

XHTML: A Bridge To The Future -- THE W3C'S RECOMMENDATION BLENDS XML AND HTML TO PRODUCE EXTENSIBLE WEB-PAGE FORMATTING. (Company Business and Marketing)

Kiely, Don

InformationWeek, 210
May 8, 2000
ISSN: 8750-6874 LANGUAGE: English RECORD TYPE: Fulltext; Abstract
WORD COUNT: 2009 LINE COUNT: 00165

... the three required DTDs.

Minimal XHTML

The following code, taken from the XHMTL proposed recommendation, is an example of a minimal XHTML 1.0 document:

<? **xml** version="1.0" encoding="UTF-8"?>

<!DOCTYPE html

PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"

"DTD/xhtml1-strict.dtd">

<html xmlns="http://www.w3.org/1999/xhtml" **xml** :lang="en" lang="en">

<head>

<title>Virtual Library</title>

</head>

<body>

Moved to vlib.org .

</body>

</html>

Some specifics...

...Empty elements must either have an end tag, or the start tag must end with />. This is sometimes called a self-terminating element.

Element and **attribute names** . **XML** is case-sensitive, and the XHTML DTDs, element, and **attribute names** must be in lower case. All **attribute values** must be quoted in single or double quotes:

Nested elements. Elements must also be properly nested, so that closing tags must be in reverse order of the opening tags.

No minimized attributes. **XML** , and therefore XHTML, does not support attribute minimization.

Script and style tags. Because any < and & characters are considered parts of tags in XHTML, any script...

16/3,K/4 (Item 4 from file: 275)
DIALOG(R)File 275:Gale Group Computer DB(TM)
(c) 2003 The Gale Group. All rts. reserv.

02339501 SUPPLIER NUMBER: 56077241 (USE FORMAT 7 OR 9 FOR FULL TEXT)
XML: Ready for Prime Time.(Technology Information)
Angel, Jonathan
Network, NA
Oct 1, 1999
ISSN: 1093-8001 LANGUAGE: English RECORD TYPE: Fulltext; Abstract
WORD COUNT: 4266 LINE COUNT: 00354

... 3>Network and Systems Management with **XML** </article...

...report>
Once elements have been defined, a DTD may also define attributes using the !ATTLIST command. This specifies an element, **names** the **attribute** to be associated with it, and then exerts con- trol over the **values** that **attribute** can have. For example, the following would associate the attribute manufacturer with the element car, allowing the former to have one of the three values...include !ENTITY declarations, which define entity references, and !NOTATION declarations, which let a parser know what to do with binary files that are not in **XML** format.
A serious and rather surprising limitation of DTDs, however, is that they do not permit datatyping-that is, constraining data to a particular format...

16/3,K/5 (Item 5 from file: 275)
DIALOG(R)File 275:Gale Group Computer DB(TM)
(c) 2003 The Gale Group. All rts. reserv.

02249947 SUPPLIER NUMBER: 53359796 (USE FORMAT 7 OR 9 FOR FULL TEXT)
SoftQuad Forges XML Trail.(XMetal authoring tool)(Software Review)(Evaluation)
Gonsalves, Antone
PC Week, 52(1)
Dec 7, 1998
DOCUMENT TYPE: Evaluation ISSN: 0740-1604 LANGUAGE: English
RECORD TYPE: Fulltext
WORD COUNT: 421 LINE COUNT: 00037

Softquad Software Inc. is one of the latest vendors to build a new tool simplifying development in **XML** , which is attracting growing corporate interest as a standard for distributing data across Web-based environments.
XMetal, scheduled to ship in March, is one of the few authoring tools that provide features to speed up the process of writing in **Extensible Markup** Language.
"There's very few **XML** /SGML (Standard Generalized ML) editors on the market right now," said Betty Harvey, an analyst at Electronic Commerce Connection Inc., of Germantown, Md., a consulting...
...which resembles a standard word processor, requires users to first choose a DTD (Document Type Definition) and configure the tool bars to insert related tag **names** , **attributes** , **attribute** **values** and other elements on demand.
"XMetal will only let you use those **XML** elements that are valid for the DTD you specify," Harvey said.
DTDs define the format codes, or tags, embedded within **XML** documents. Each industry--health care, insurance and so on--uses different DTDs.
XMetal has an extensible object management system, called Resource Manager, for storing boilerplate...

...Microsoft Corp.'s COM (Component Object Model), developers can use any tool that supports COM to build wizards and add them to the XMetal editor.

XML is a subset of SGML, a text-based language for describing the content and structure of digital documents.

The wizards can be used to build customized functions within the **XML** documents. The new product allows developers to save SGML documents as **XML** , providing an easy migration path.

XMetal supports HTML 4.0 elements, such as forms, tables and links, as well as other standards approved by the...

16/3,K/6 (Item 6 from file: 275)
DIALOG(R)File 275:Gale Group Computer DB(TM)

02187478 SUPPLIER NUMBER: 20813318 (USE FORMAT 7 OR 9 FOR FULL TEXT)
Object-Based Web Design.(browser-based document object model) (PC Tech) (Technology Tutorial)(Tutorial)(Column)
Stanek, William Robert
PC Magazine, v17, n14, p295(5)
July, 1998
DOCUMENT TYPE: Tutorial Column ISSN: 0888-8507 LANGUAGE: English
RECORD TYPE: Fulltext; Abstract
WORD COUNT: 4358 LINE COUNT: 00347

... can be invoked to update the element. In this preliminary definition, there are three basic methods: addAttribute, getAttribute, and removeAttribute. The addAttribute method sets a **value** for a **named attribute** ; getAttribute method returns a reference to a **named attribute** so that you can examine its values; and removeAttribute does just that, which effectively sets some attributes back to their default value.

Language-specific Bindings...

...are the language-specific bindings for the document object model, which map language-specific calls and data to language-neutral calls and data. The core, **XML** , and HTML object models have different sets of bindings. Because those bindings are language-specific, they must be implemented in each programming language separately. That...

16/3,K/7 (Item 7 from file: 275)
DIALOG(R)File 275:Gale Group Computer DB(TM)

02180236 SUPPLIER NUMBER: 20652651 (USE FORMAT 7 OR 9 FOR FULL TEXT)
Structuring Data with XML. (Internet/Web/Online Service Information)(Column)
Stanek, William Robert
PC Magazine, v17, n10, p229(1)
May 26, 1998
DOCUMENT TYPE: Column ISSN: 0888-8507 LANGUAGE: English
RECORD TYPE: Fulltext; Abstract
WORD COUNT: 4490 LINE COUNT: 00381

... transaction as incomplete and inform the reader that the current transaction is pending.

The final attribute for the purchase element is an enumerated type. The **attribute** has four **name** tokens that define the valid **values** for the **attribute** . From these tokens, you can determine that purchase transactions are defined as walk-in orders, Web-related orders, phone orders or mail orders. The default **value** for the **attribute** is "walkin." Thus if no value is specified, the purchase transaction is assumed to be a walk-in order.

Completing the DTD

After defining the...

...add some additional information to the document, such as your company name or copyright information. Although you could insert this information directly into the document, **XML** does provide a mechanism that makes it easier to maintain documents over time. This mechanism is called an entity.

In its basic form, an entity...

...entities allow text and files to be substituted into a document, they can be used to replace values when a document is displayed.

Unlike HTML, **XML** allows you to define your own entities. Let's say your company name is CPC Enterprises Ltd., and you want to add this information to...

16/3,K/8 (Item 1 from file: 621)
DIALOG(R)File 621:Gale Group New Prod.Annou.(R)

01769741 Supplier Number: 53364759 (USE FORMAT 7 FOR FULLTEXT)
Veo Systems Announces Support For Sun Microsystems' New Java Development Kit With Java-Based Parser; Veo Systems Leads Integration of XML and Java.
Business Wire, p0011
Dec 8, 1998
Language: English Record Type: Fulltext
Document Type: Newswire; Trade
Word Count: 390

Due to ship in Q1 of 1999, Veo Systems' **XML** parser will bridge the gap between **XML** and Java by being the first parser to support the Schema for Object Oriented **XML** (SOX) submission to the W3C **XML** -Schema working group. The Veo parser will combine unprecedented speed with **XML** DTD and SOX validation.

SOX extends **XML** for e-commerce and distributed computing by adding strong typing, inheritance, global **name** spaces, and legal **attribute values** . These extensions enable validity checking and facilitate mapping a document to other documents as well as to the rich data schemas used in databases and object oriented software applications.

"Veo is leading the way to couple the power of **XML** and Java. Veo Systems' parser will bridge the gap between **XML** and Java and by giving developers a familiar typed programming model for manipulating **XML** documents and make **XML** more powerful for the Java developer," stated Dr. Jay Tenenbaum, chairman and chief scientist of Veo Systems. "We are looking forward to working with our customers who deploy Sun's JDK in building innovative e-commerce and business integration solutions."

Veo Systems is currently accepting applicants for the **XML** parser early access program. Anyone interested in participating should visit Veo's website at www.veosystems.com/ **xml** /parser/parser.html

About Veo Systems, Inc.

Veo Systems is the leading supplier of products and services to enable open commerce networks. Using Veo's solutions, companies can significantly lower the economics of business-to-business integration by exchanging information using self-describing **XML** business documents that both people and computers can readily understand.

Veo's world-class technology team pioneered the application of **XML** to electronic commerce. Veo Systems is located in Mountain View, Calif. and can be reached at 650/988-7244 or via the Internet at http...

16/3,K/9 (Item 1 from file: 16)
DIALOG(R)File 16:Gale Group PROMT(R)

05350421 Supplier Number: 48138797 (USE FORMAT 7 FOR FULLTEXT)
Extend The Web: An XML Primer
Powell, Thomas A.
InternetWeek, p47
Nov 24, 1997
Language: English Record Type: Fulltext
Document Type: Newsletter; Trade
Word Count: 3094

... be a boon. A simplified version of SGML for the Web would be one logical way to address the limitations of HTML.

The Rules Of **XML**

This is where **XML** comes in. Its design goal was to be a subset of SGML useful for the Web. Writing **XML** sounds like a daunting task, but this isn't necessarily true. Unlike HTML, miscoded documents in **XML** won't render at all. For those Webmasters considering how many of their HTML documents would render if that language had such strict rule enforcement, you can stop sweating. The rules of basic **XML** are easy. Suppose you need to define some tags to represent an invoice in **XML** --merely create a document as shown in the green chart below:

Naming a tag <ENTRY> instead of <ITEM> is completely up to the document author. Choose any element and **attribute names** that represent the domain being modeled. Does **XML** have any more rules? Yes, but they are few, and only relate to syntax.

The first rule is that, just like well-written HTML, all ...tags must be properly nested and must match. There also must be an enclosing element for the whole document.

The second rule is that all **attribute values** must be quoted. In HTML this is good authoring practice but is only required for values that contain spaces.

In **XML** , BLASTOFFCOUNT="10"> is correct. <BLASTOFF COUNT=10> is not correct.

The final rule is that all elements with empty content must be self-identifying by...

...instead of the familiar ">". An empty element is one that does not require a closing tag, like the HTML elements
, <HR> and <IMG>. In **XML** these would be
, <HR/> and <IMG SRC="test.gif"/>. Why is this last rule needed? Because **XML** documents may not have a formal DTD associated with them. Lacking a DTD, there is no way for a parser to know if a tag like
 is empty or requires a matching </BR> tag later in the document. For parsing efficiency, **XML** needs a syntactic signal to identify empty tags.

An **XML** document made according to these rules is a "well-formed" document. SGML purists may find this notion eccentric and somewhat troubling. However, just because a...

16/3,K/10 (Item 2 from file: 16)

05283580 Supplier Number: 48047330 (USE FORMAT 7 FOR FULLTEXT)

W3C founding member Murray Maloney discusses XML for meta data

Radosevich, Lynda

InfoWorld, p73
Oct 13, 1997
Language: English Record Type: Fulltext
Document Type: Magazine/Journal; Trade
Word Count: 685

InfoWorld: Is one of the potential uses of **XML** to enable more sophisticated searching than is available today?

Maloney: This is one of the most obvious and compelling benefits **XML** could provide.

InfoWorld: To enable searches on specific categories, say on "Welsh corgis," won't vertical industry segments have to agree on tags for **XML** -based markup languages such as the Dog Breeder's Markup Language (DBML)?

Maloney: That is darned close to being completely accurate. The SGML [Standard Generalized Markup Language]/ **XML** pedants would tell you that you don't have to agree on a set of tags [element types], but could instead agree on a set of **attribute names** and **values** . In fact, the use of attributes to provide deeper meaning is a characteristic of "architectural forms." The good news is that there are quite a...

...already deployed. These tools are equipped to handle any number of markup languages.

In general, Web tools will have to be equipped to read an **XML** document, present it according to the rules specified by a style sheet

(which may be Cascading Style [redacted]eets, Extensible Style Sheets, [redacted] some other style...

16/3,K/11 (Item 1 from file: 148)
DIALOG(R)File 148:Gale Group Trade & Industry DB

09790965 SUPPLIER NUMBER: 19868163 (USE FORMAT 7 OR 9 FOR FULL TEXT)
Speak my language: W3C founding member Murray Maloney discusses XML for meta data. (World Wide Web Consortium, Extensible Markup Language) (Technology Information) (Interview)
Radosevich, Lynda
InfoWorld, v19, n41, p73(2)
Oct 13, 1997
DOCUMENT TYPE: Interview ISSN: 0199-6649 LANGUAGE: English
RECORD TYPE: Fulltext; Abstract
WORD COUNT: 724 LINE COUNT: 00059

InfoWorld: Is one of the potential uses of **XML** to enable more sophisticated searching than is available today?

Maloney: This is one of the most obvious and compelling benefits **XML** could provide.

InfoWorld: To enable searches on specific categories, say on "Welsh corgis," won't vertical industry segments have to agree on tags for **XML** -based markup languages such as the Dog Breeder's Markup Language (DBML)?

Maloney: That is darned close to being completely accurate. The SGML (Standard Generalized Markup Language)/ **XML** pedants would tell you that you don't have to agree on a set of tags (element types), but could instead agree on a set of **attribute names** and **values** . In fact, the use of attributes to provide deeper meaning is a characteristic of "architectural forms." The good news is that there are quite a...

...already deployed. These tools are equipped to handle any number of markup languages.

In general, Web tools will have to be equipped to read an **XML** document, present it according to the rules specified by a style sheet (which may be Cascading Style Sheets, Extensible Style Sheets, or some other style...

16/3,K/12 (Item 1 from file: 15)
DIALOG(R)File 15:ABI/Inform(R)

01991081 46463881
W3C readies XHTML for approval
Piven, Joshua
Computer Technology Review v19n10 PP: 4 Oct 1999
ISSN: 0278-9647 JRNL CODE: GBAMN
WORD COUNT: 440

...TEXT: phase, the next-to-last step before XHTML becomes an official Recommendation. The next-- generation Web standard-for all purposes a merger of HTML and **XML** -should be approved by the time you read this.

XHTML is the most significant change to the language of the Web since HTML 4.0...

... approved in 1997. XHTML is a family of document types and modules that extend the functionality of HTML 4.0 and has its roots in **XML** . Documents based on XHTML are designed to work with the new wave of Webenabled devices, including cell phones and PDAs.

In brief, XHTML documents conform to **XML** and can be viewed, edited, and validated with standard **XML** tools. However, XHTML documents can also be viewed by existing HTML 4.0-compliant browsers and other user agents. XHTML documents can also run processes (scripts and applets) that are based on

the HTML Document Object Model (DOM) or the **XML** DOM.

XHTML contains several important syntax changes from HTML. Since **XML** is case-- sensitive, XHTML documents must use lower case for all HTML element and **attribute names** . Also, XHTML is strict in its interpretation of tags; this means that all elements must either have closing tags or be written in a special...

... creates a paragraph break, is often used without its closing partner </p>. This is unacceptable in XHTML. Other changes include the need for quotes in **attribute values** . For example, the **attribute** tag <table rows="3"> is correct; <table rows=3> is incorrect.

In creating XHTML documents, authors should note that pages can be labeled as text/html, text/ **xml** , or application/ **xml** . When labeled as text/html, however, documents that don't follow standard HTML Compatibility Guidelines will almost certainly fail to be processed, according to XHTML...

... will be carried out on alternate platforms-that is, non PC-based applications. New protocols like WAP, which target small-footprint microbrowsers, are already implementing **XML** . However, Web content must be written specifically in WML, WAP's markup language, in order to be processed quickly and efficiently. XHTML will give page...

16/3,K/13 (Item 1 from file: 647)
DIALOG(R)File 647:CMP Computer Fulltext

01227076 CMP ACCESSION NUMBER: NWC20001113S0026
XHTML: Crossroads of HTML and XML - Sloppy coding habits just won't work with XHTML. We'll help clean up your language.
Ahmad Abualsamid
NETWORK COMPUTING, 2000, n 1122, PG155
PUBLICATION DATE: 001113
JOURNAL CODE: NWC LANGUAGE: English
RECORD TYPE: Fulltext
SECTION HEADING: Workshop - Development
WORD COUNT: 1973

... form elements. To make sure your code works with today's browsers and with existing scripts, you can use both a name and an ID **attribute** .

- **Attribute values** must be quoted, and no minimization is allowed. A common practice among HTML coders is to leave the quotes out when specifying values for elements...

...in XHTML. Some attributes, such as "checked," could be minimized when using several browsers. This also is not valid. You can't have a dangling **attribute** :

< input id="acheckbox" **name** ="acheckbox" checked /> is incorrect.

< input id="acheckbox" name="acheckbox" checked= "checked" /> is correct.

- XHTML documents have some mandatory elements. You no longer can have documentsbody >

</body

</html

Rules of **XML**

Another key difference is that XHTML documents must conform to **XML** rules. Here are the pertinent **XML** rules for XHMTL developers (for more information, see www. **xml** .com):

- All **XML** documents are well-formed by definition. A well-formed document adheres to the **XML** structure but does not follow a certain DTD. A document following a certain DTD is called valid. In XHTML, the DTD is that of HTML...

16/3,K/14 (Item 2 from file: 647)
DIALOG(R)File 647:CMP Computer Fulltext

01215262 CMP ACCESSION NUMBER: IWK20000508S0072
XHTML: A Bridge To The Future - THE W3C'S RECOMMENDATION BLENDS XML AND HTML TO PRODUCE EXTENSIBLE WEB-PAGE FORMATTING
DON KIELY
INFORMATIONWEEK, 2000, n 785, PG210
PUBLICATION DATE: 000508
JOURNAL CODE: IWK LANGUAGE: English
RECORD TYPE: Fulltext
SECTION HEADING: Application Development
WORD COUNT: 1903

... the three required DTDs.

Minimal XHTML

The following code, taken from the XHMTL proposed recommendation, is an example of a minimal XHTML 1.0 document:

<? **xml** version="1.0" encoding="UTF-8"?>

PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"

"DTD/xhtml1-strict.dtd">

<html xmlns="http://www.w3.org/1999/xhtml" **xml** :lang="en" lang="en">

<head>

<title>Virtual Library</title

</head

<body>

Moved to <a href="http://vlib.org/">vlib.org </a.

</body

</html

Some specifics...

...Empty elements must either have an end tag, or the start tag must end with />. This is sometimes called a self- terminating element.

Element and **attribute** names . **XML** is case-sensitive, and the XHTML DTDs, element, and **attribute** **names** must be in lower case. All **attribute** **values** must be quoted in single or double quotes:

Nested elements. Elements must also be properly nested, so that closing tags must be in reverse order of the opening tags.

No minimized attributes. **XML** , and therefore XHTML, does not support attribute minimization.

Script and style tags. Because many < and & characters are considered parts of tags in XHTML, any script...

16/3,K/15 (Item 3 from file: 647)
DIALOG(R)File 647:CMP Computer Fulltext

01146357 CMP ACCESSION NUMBER: INW19971124S0090
Extend The Web: An XML Primer
Thomas A. Powell
INTERNETWEEK, 1997, n 691, PG47
PUBLICATION DATE: 971124
JOURNAL CODE: INW LANGUAGE: English
RECORD TYPE: Fulltext
SECTION HEADING: Reviews
WORD COUNT: 3059

... be a boon. A simplified version of SGML for the Web would be one logical way to address the limitations of HTML.

The Rules Of **XML**

This is where **XML** comes in. Its design goal was to be a subset of SGML useful for the Web. Writing **XML** sounds like a daunting task, but this isn't necessarily true. Unlike HTML, miscoded documents in **XML** won't render at all. For those Webmasters considering how many of their HTML documents would render if that language had such strict rule enforcement, you can stop sweating. The rules of basic **XML** are easy. Suppose you need to define some tags to represent an invoice in **XML** -merely create a document as shown in the green chart below:

Naming a tag <ENTRY>instead of <ITEM>is completely up to the document author. Choose any element and **attribute names** that represent the domain being modeled. Does **XML** have any more rules? Yes, but they are few, and only relate to syntax.

The first rule is that, just like well-written HTML, all ...tags must be properly nested and must match. There also must be an enclosing element for the whole document.

The second rule is that all **attribute values** must be quoted. In HTML this is good authoring practice but is only required for values that contain spaces.

In **XML** , BLASTOFFCOUNT=

>"10" is correct. <BLASTOFF COUNT&EQUALS;10>is not correct.

>>The final rule is that all elements with empty content must be self-identifying...

...instead of the familiar "". An empty element is one that does not require a closing tag, like the HTML elements
, <HR>and <IMG>. In **XML** these would be
, <HR/>and <IMG SRC&EQUALS;"test.gif"/>. Why is this last rule needed? Because **XML** documents may not have a formal DTD associated with them. Lacking a DTD, there is no way for a parser to know if a tag like
is empty or requires a matching</BRtag later in the document. For parsing efficiency, **XML** needs a syntactic signal to identify empty tags.

An **XML** document made accordin